Method for Identifying a Transient Acoustic Scene, Application of Said Method, and a Hearing Device

This invention relates to a method for identifying a transient acoustic scene, an application of said method in conjunction with hearing devices, as well as a hearing device.

Modern-day hearing aids, when employing different audiophonic programs -- typically two to a maximum of three such hearing programs -- permit their adaptation to varying acoustic environments or scenes. The idea is to optimize the effectiveness of the hearing aid for its user in all situations.

The hearing program can be selected either via a remote control or by means of a selector switch on the hearing aid itself. For many users, however, having to switch program settings is a nuisance, or difficult, or even impossible. Nor is it always easy even for experienced wearers of hearing aids to determine at what point in time which program is most comfortable and offers optimal speech discrimination. An automatic recognition of the acoustic scene and corresponding automatic switching of the program setting in the hearing aid is therefore desirable.

There exist several different approaches to the automatic classification of acoustic surroundings. All of the methods concerned involve the extraction of different characteristics from the input signal which may be derived from one or several microphones in the hearing aid. Based on these characteristics, a pattern-recognition device employing a particular algorithm makes a determination as to the attribution of

the analyzed signal to a specific acoustic environment. These various existing methods differ from one another both in terms of the characteristics on the basis of which they define the acoustic scene (signal analysis) and with regard to the pattern-recognition device which serves to classify these characteristics (signal identification).

For the extraction of characteristics in audio signals, J.M. Kates in his article titled "Classification of Background Noises for Hearing-Aid Applications" (1995, Journal of the Acoustical Society of America 97(1), pp 461-469), suggested an analysis of time-related sound-level fluctuations and of the sound spectrum. On its part, the European patent EP-B1-0 732 036 proposed an analysis of the amplitude histogram for obtaining the same result. Finally, the extraction of characteristics has been investigated and implemented based on an analysis of different modulation frequencies. In this connection, reference is made to the two papers by Ostendorf et al titled "Empirical Classification of Different Acoustic Signals and of Speech by Means of a Modulation-Frequency Analysis" (1997, DAGA 97, pp 608-609), and "Classification of Acoustic Signals Based on the Analysis of Modulation Spectra for Application in Digital Hearing Aids" (1998, DAGA 98, pp 402-403). A similar approach is described in an article by Edwards et al titled "Signal-processing algorithms for a new software-based, digital hearing device" (1998, The Hearing Journal 51, pp 44-52). Other possible characteristics include the sound-level transmission itself or the zero-passage rate as described for instance in the article by H.L. Hirsch, titled "Statistical Signal Characterization" (Artech House 1992). It is evident that the characteristics used to

date for the analysis of audio signals are strictly based on system-specific parameters.

It is fundamentally possible to use prior-art pattern identification methods for sound classification purposes. Particularly suitable pattern-recognition systems are the so-called ranging devices, Bayes classifiers, fuzzy-logic systems and neural networks. Details of the first two of the methods mentioned are contained in the publication titled "Pattern Classification and Scene Analysis" by Richard O. Duda and Peter E. Hart (John Wiley & Sons, 1973). For information on neural networks, reference is made to the treatise by Christopher M. Bishop, titled "Neural Networks for Pattern Recognition" (1995, Oxford University Press). Reference is also made to the following publications: Ostendorf et al, "Classification of Acoustic Signals Based on the Analysis of Modulation Spectra for Application in Digital Hearing Aids" (Zeitschrift für Audiologie (Journal of Audiology), pp 148-150); F. Feldbusch, "Sound Recognition Using Neural Networks" (1998, Journal of Audiology, pp 30-36); European patent application, publication number EP-A1-0 814 636; and US patent, publication number US-5.604.812. Yet all of the pattern-recognition methods mentioned are deficient in one respect in that they merely model static properties of the sound categories of interest.

One shortcoming of these earlier sound-classification methods, involving characteristics extraction and pattern recognition, lies in the fact that, although unambiguous and solid identification of voice signals is basically possible, a number of different acoustic situations cannot be satisfactorily classified, or not at all. While these earlier methods permit a distinction between pure voice or speech signals and "non-speech" sounds,

meaning all other acoustic surroundings, that is not enough for selecting an optimal hearing program for a transient acoustic situation. It follows that the number of possible hearing programs is limited to those two automatically recognizable acoustic situations or the hearing-aid wearer himself has to recognize the acoustic situations that are not covered and manually select the appropriate hearing program.

It is therefore the objective of this invention to introduce first of all a method for identifying a transient acoustic scene which compared to prior-art methods is substantially more reliable and more precise.

This is accomplished by the measures specified in claim 1. Additional claims specify advantageous enhancements of the invention, an application of the method, as well as a hearing device.

The invention is based on an extraction of signal characteristics, a subsequent separation of different sound-sources as well as an identification of different sounds. In lieu of or in addition to system-specific characteristics, auditory characteristics are taken into account in the signal analysis for the extraction of characteristic features. These auditory characteristics are identified by means of Auditory Scene Analysis (ASA) techniques. In another form of implementation of the method per this invention, the characteristics are subjected to a context-free or a context-sensitive grouping process by applying the gestalt principle. The actual identification and classification of the audio signals derived from the extracted characteristics is preferably performed using Hidden Markov Models (HMM). One advantage of this invention is the fact that it allows for

a larger number of identifiable sound categories and thus a greater number of hearing programs which results in enhanced sound classification and correspondingly greater comfort for the user of the hearing device.

The following will explain this invention in more detail by way of an example with reference to a drawing. The only figure is a functional block diagram of a hearing device in which the method per this invention has been implemented.

In the figure, the reference number 1 designates a hearing device. For the purpose of the following description, the term "hearing device" is intended to include hearing aids as used to compensate for the hearing impairment of a person, but also all other acoustic communication systems such as radio transceivers and the like.

The hearing device 1 incorporates in conventional fashion two electro-acoustic converters 2a, 2b and 6, these being one or several microphones 2a, 2b and a speaker 6, also referred to as a receiver. A main component of a hearing device 1 is a transmission unit 4 in which, in the case of a hearing aid, signal modification takes place in adaptation to the requirements of the user of the hearing device 1. However, the operations performed in the transmission unit 4 are not only a function of the nature of a specific purpose of the hearing device 1 but are also, and especially, a function of the momentary acoustic scene. There have already been hearing aids on the market where the wearer can manually switch between different hearing programs tailored to specific acoustic situations. There also exist hearing aids capable of automatically

recognizing the acoustic scene. In that connection, reference is again made to the European patents EP-B!-0 732 036 and EP-A1-0 814 636 and to the US patent 5.604.812, as well as to the "Claro Autoselect" brochure by Phonak Hearing Systems (28148 (GB) /0300, 1999).

In addition to the aforementioned components such as microphones 2a, 2b, the transmission unit 4 and the receiver 6, the hearing device 1 contains a signal analyzer 7 and a signal identifier 8. If the hearing device 1 is based on digital technology, one or several analog-to-digital converters 3a, 3b are interpolated between the microphones 2a, 2b and the transmission unit 4 and one digital-to-analog converter 5 is provided between the transmission unit 4 and the receiver 6. While a digital implementation of this invention is preferred, it should be equally possible to use analog components throughout. In that case, of course, the converters 3a, 3b and 5 are not needed.

The signal analyzer 7 receives the same input signal as the transmission unit 4. The signal identifier 8, which is connected to the output of the signal analyzer 7, connects at the other end to the transmission unit 4 and to a control unit 9.

A training unit 10 serves to establish in off-line operation the parameters required in the signal identifier 8 for the classification process.

By means of a user input unit 11, the user can override the settings of the transmission

unit 4 and the control unit 9 as established by the signal analyzer 7 and the signal identifier 8.

The method according to this invention is explained as follows:

It is essentially based on the extraction of characteristic features from an acoustic signal during an extraction phase, whereby, in lieu of or in addition to the system-specific characteristics -- such as the above-mentioned zero-passage rates, time-related sound-level fluctuations, different modulation frequencies, the sound level itself, the spectral peak, the amplitude distribution etc. -- auditory characteristics as well are employed. These auditory characteristics are determined by means of an Auditory Scene Analysis (ASA) and include in particular the volume, the spectral pattern (timbre), the harmonic structure (pitch), common build-up and decay times (on-/offsets), coherent amplitude modulations, coherent frequency modulations, coherent frequency transitions, binaural effects etc. Detailed descriptions of Auditory Scene Analysis can be found for instance in the articles by A. Bregman, "Auditory Scene Analysis" (MIT Press, 1990) and W.A. Yost, "Fundamentals of Hearing - An Introduction" (Academic Press, 1977). The individual auditory characteristics are described, inter alia, by A. Yost and S. Sheft in "Auditory Perception" (published in "Human Psychophysics" by W.A. Yost, A.N. Popper and R.R. Fay, Springer 1993), by W.M. Hartmann in "Pitch, Periodicity, and Auditory Organization" (Journal of the Acoustical Society of America, 100 (6), pp 3491-3502, 1996), and by D.K. Mellinger and B.M. Mont-Reynaud in "Scene Analysis" (published

in "Auditory Computation" by H.L. Hawkins, T.A. McMullen, A.N. Popper and R.R. Fay, Springer 1996).

In this context, an example of the use of auditory characteristics in signal analysis is the characterization of the tonality of the acoustic signal by analyzing the harmonic structure, which is particularly useful in the identification of tonal signals such as speech and music.

Another form of implementation of the method according to this invention additionally provides for a grouping of the characteristics in the signal analyzer 7 by means of gestalt analysis. This process applies the principles of the gestalt theory, by which such qualitative properties as continuity, proximity, similarity, common destiny, unity, good constancy and others are examined, to the auditory and perhaps system-specific characteristics for the creation of auditory objects. This grouping -- and, for that matter, the extraction of characteristics in the extraction phase -- can take place in context-free fashion, i.e. without any enhancement by additional knowledge (so-called "primitive" grouping), or in context-sensitive fashion in the sense of human auditory perception employing additional information or hypotheses regarding the signal content (so-called "design-based" grouping). This means that the contextual grouping is adapted to any given acoustic situation. For a detailed explanation of the principles of the gestalt theory and of the grouping process employing gestalt analysis, substitutional reference is made to the publications titled "Perception Psychology" by E.B. Goldstein (Spektrum Akademischer Verlag, 1997), "Neural Fundamentals of Gestalt Perception" by A.K.

Engel and W. Singer (Spektrum der Wissenschaft, 1998, pp 66-73), and "Auditory Scene Analysis" by A. Bregman (MIT Press, 1990).

The advantage of applying this grouping process lies in the fact that it allows further differentiation of the characteristics of the input signals. In particular, signal segments are identifiable which originate in different sound-sources. The extracted characteristics can thus be mapped to specific individual sound sources, providing additional information on these sources and, hence, on the current, transient auditory scene.

The second aspect of the method according to this invention as described here relates to pattern recognition, i.e. the signal identification that takes place during the identification phase. The preferred form of implementation of the method per this invention employs the Hidden Markov Model (HMM) method in the signal identifier 8 for the automatic classification of the acoustic scene. This also permits the use of time changes of the computed characteristics for the classification process. Accordingly, it is possible to also take into account dynamic and not only static properties of the surrounding situation and of the sound categories. Equally possible is a combination of HMMs with other classifiers such as multi-stage recognition processes for identifying the acoustic scene.

The output signal of the signal identifier 8 thus contains information on the nature of the acoustic surroundings (the acoustic situation or scene). That information is fed to the transmission unit 4 which selects the program, or set of parameters, best suited to the

transmission of the acoustic scene discerned. At the same time, the information gathered in the signal identifier 8 is fed to the control unit 9 for further actions whereby, depending on the situation, any given function, such as an acoustic signal, can be triggered.

If the identification phase involves Hidden Markov Models, it will require a complex process for establishing the parameters needed for the classification. This parameter ascertainment is therefore best done in the off-line mode, individually for each category or class at a time. The actual identification of various acoustic scenes requires very little memory space and computational capacity. It is therefore recommended that a training unit 10 be provided which has enough computing power for parameter determination and which can be connected via appropriate means to the hearing device 1 for data transfer purposes. The connecting means mentioned may be simple wires with suitable plugs.

The method according to this invention thus makes it possible to select from among numerous available settings and automatically pollable actions the one best suited without the need for the user of the device to make the selection. This makes the device significantly more comfortable for the user since upon the recognition of a new acoustic scene it promptly and automatically selects the right program or function in the hearing device 1.

The users of hearing devices often want to switch off the automatic recognition of the acoustic scene and corresponding automatic program selection, described above. For

this purpose a user input unit 11 is provided by means of which it is possible to override the automatic response or program selection. The user input unit 11 may be in the form of a switch on the hearing device 1 or a remote control which the user can operate. There are also other options which offer themselves, for instance a voice-activated user input device.